FEEDBACK-RESISTANT PYRUVATE CARBOXYLASE GENE FROM CORYNEBACTERIUM

Inventor: Paul D. Hanke

CROSS-REFERENCE TO RELATED APPLICATIONS

This application claims the benefit of the filing date of U.S. Provisional Application No. 60/239,913, filed on October 13, 2000, which is incorporated herein by reference in its entirety.

BACKGROUND OF THE INVENTION

Field of the Invention

[0001] The present invention relates to a mutated pyruvate carboxylase gene from *Corynebacterium*. The mutant pyruvate carboxylase gene encodes a pyruvate carboxylase enzyme which is resistant to feedback inhibition from aspartic acid. The present invention also relates to a method of replacing the wild-type pyruvate carboxylase gene in *Corynebacterium* with this feedback-resistant pyruvate carboxylase gene. The present invention further relates to methods of the production of amino acids, preferably lysine, comprising the use of this mutant pyruvate carboxylase enzyme in microorganisms.

Background Art

[0002] Pyruvate carboxylase is an important biotin-containing enzyme found in a variety of plants and animals, as well as some groups of bacteria (Modak, H.V. and Kelly, D.J., Microbiology 141:2619-2628 (1995)). In the presence of adenosine triphosphate (ATP) and magnesium ions, pyruvate carboxylase catalyzes the two-step carboxylation of pyruvate to form oxaloacetate, as shown in the equations below:

$$MgATP + HCO_3^- + ENZ\text{-}biotin \xrightarrow{Mg^{2+}acetyl\text{-}CoA} MgADP + Pi + ENZ\text{-}biotin\text{-}CO_2 \quad (1)$$

$$ENZ\text{-}biotin\text{-}CO_2 + Pyruvate \longrightarrow ENZ\text{-}biotin + oxaloacetate \quad (2)$$

[0003] In reaction (1) the ATP-dependent biotin carboxylase domain carboxylates a biotin prosthetic group linked to a specific lysine residue in the biotin-carboxyl-carrier protein (BCCP) domain. Acetyl-coenzyme A activates reaction (1) by increasing the rate of bicarbonate-dependent ATP cleavage. In reaction (2), the BCCP domain donates the CO_2 to pyruvate in a reaction catalyzed by the transcarboxylase domain (Attwood, P.V., *Int. J. Biochem. Cell. Biol. 27*:231-249 (1995)).

[0004] In bacteria such as *Corynebacterium glutamicum*, pyruvate carboxylase is utilized during carbohydrate metabolism to form oxaloacetate, which is in turn used in the biosynthesis of amino acids, particularly L-lysine and L-glutamate. Furthermore, in response to a cell's metabolic needs and internal environment, the activity of pyruvate carboxylase is subject to both positive and negative feedback mechanisms, where the enzyme is activated by acetyl-CoA, and inhibited by aspartic acid. Based on its role in the pathway of amino acid synthesis, and its ability to be regulated, pyruvate carboxylase plays a vital role in the synthesis of amino acids.

[0005] Bacteria such as *C. glutamicum* and *E. coli* are widely used in industry for the production of amino acids such as L-glutamate and L-lysine. Because of the central importance of pyruvate carboxylase in the production of amino acids, particularly L-glutamate and L-lysine, the exploitation of pyruvate carboxylase to increase amino acid production is of great interest in an industrial setting. Thus, promoting the positive feedback mechanism of pyruvate carboxylase, or

inhibiting its negative feedback mechanism, in *C. glutamicum* or could augment amino acid production on an industrial scale.

BRIEF SUMMARY OF THE INVENTION

[0006] One aspect of the present invention relates to a nucleic acid molecule comprising a nucleotide sequence which codes for a pyruvate carboxylase of SEQ ID NO:19, wherein this pyruvate carboxylase contains at least one mutation which desensitizes the pyruvate carboxylase to feedback inhibition by aspartic acid.

[0007] Another aspect of the present invention provides methods for using the nucleic acid of SEQ ID NO:1 or SEQ ID NO:3, which encodes the amino acid sequence of a mutant pyruvate carboxylase. Such uses include the replacement of the wild-type pyruvate carboxylase with the feedback-resistant pyruvate carboxylase, and the production of amino acids. An additional aspect of the present invention provides a polypeptide comprising the amino acid sequence of SEQ ID NO:2 or SEQ ID NO:4. Still another aspect of the present invention provides a polypeptide comprising the amino acid sequence selected from the group comprising SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:12 SEQ ID NO:14, SEQ ID NO:16 and SEQ ID NO:18.

[0008] Another aspect of the present invention also relates to a nucleic acid molecule comprising a nucleotide sequence which encodes the amino acid sequence of SEQ ID NO:2, SEQ ID NO:4, or the amino acid sequence encoded by the DNA contained in Deposit Number NRRL B-11474. Another aspect of the present invention further relates to a nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:1 and SEQ ID NO:3.

BRIEF DESCRIPTION OF THE DRAWINGS/FIGURES

[0009] Figures 1A-1E show the full-length nucleotide sequence (SEQ ID NO:1) encoding the amino acid sequence of feedback-resistant pyruvate carboxylase, and the corresponding amino acid sequence (SEQ ID NO:2).

[0010] Figure 2 shows the comparison of amino acid sequences between the wild-type pyruvate carboxylase, isolated from ATCC21253, and the feedback-resistant pyruvate carboxylase (SEQ ID NO:2), isolated from Deposit Number NRRL B-11474.

[0011] Figures 3A-3B show the full-length nucleotide sequence (SEQ ID NO:3) encoding the amino acid sequence of feedback-resistant pyruvate carboxylase.

[0012] Figure 4 shows the effects of various substrate concentrations on the pyruvate carboxylase activity in *C. glutamicum* ATCC 21253 and NRRL B-11474.

[0013] Figure 5 shows the effects of aspartate concentration on the activity of pyruvate carboxylase in *C. glutamicum* ATCC21253 and NRRL B-11474.

[0014] Figure 6 shows the effects of acetyl-CoA concentration on the activity of pyruvate carboxylase in *C. glutamicum* ATCC21253 and NRRL B-11474.

DETAILED DESCRIPTION OF THE INVENTION

[0015] The present invention relates to variations of the polypeptide comprising the amino acid sequence which codes for the pyruvate carboxylase as shown in SEQ ID NO:19. Preferably, the variations of pyruvate carboxylase enzyme in the present invention contain at least one mutation which desensitizes the pyruvate carboxylase to feedback inhibition by aspartic acid. Such mutations may include deletions, insertions, inversions, repeats, and type substitutions. More preferably, the amino acid sequence mutation which desensitizes the wild-type pyruvate carboxylase enzyme (SEQ ID NO:19) to feedback inhibition comprises at least

one substitution selected from the group consisting of (a) methionine at position 1 being replaced with a valine, (b) glutamic acid at position 153 being replaced with an aspartic acid, (c) alanine at position 182 being replaced with a serine, (d) alanine at position 206 being replaced with a serine, (e) histidine at position 227 being replaced with an arginine, (f) alanine at position 452 being replaced with a glycine, and (g) aspartic acid at position 1120 being replaced with a glutamic acid. Still more preferably, the variation of the polypeptide encoded by the amino acid sequence of SEQ ID NO:19 contains more than one of the above-mentioned mutations. Most preferably, the variation of the polypeptide encoded by the amino acid sequence of SEQ ID NO:19 contains all of the above-mentioned mutations. As one of ordinary skill in the art would appreciate, the numbering of amino acid residues of a protein as used herein, begins at the amino terminus (N-terminus) and proceeds towards the carboxy terminus (C-terminus), such that the first amino acid at the N-terminus is position 1.

[0016] An embodiment of the present invention relates to an isolated or purified nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of: (a) a nucleotide sequence which encodes the amino acid sequence of SEQ ID NO:2, (b) a nucleotide sequence which encodes the amino acid sequence of SEQ ID NO:4, (c) a nucleotide sequence encoding the amino acid sequence encoded by the DNA contained in Deposit Number NRRL B-11474 or (d) a nucleotide sequence complementary to any of the nucleotide sequences in (a), (b), or (c).

[0017] Further embodiments of the invention include isolated nucleic acid molecules that comprise a polynucleotide having a nucleotide sequence at least 90% identical, and more preferably at least 95%, 97%, 98%, 99% or 100% identical, to any of the nucleotide sequences in (a), (b), (c) or (d) above, or a polynucleotide which hybridizes under stringent hybridization conditions to a polynucleotide having a nucleotide sequence identical to a nucleotide sequence in (a), (b), (c) or (d) above. However, the polynucleotide which hybridizes does

not hybridize under stringent hybridization conditions to a polynucleotide having a nucleotide sequence consisting of only A residues or of only T residues.

[0018] Another aspect of the invention is directed to nucleic acid molecules at least 90%, 95%, 97%, 98% or 99% identical to the nucleic acid sequence shown in Figure 1 (SEQ ID NO:1), Figure 3 (SEQ ID NO:3) or to the nucleic acid sequence of the deposited DNA (NRRL B-30293, deposited May 30, 2000).

[0019] A further aspect of the invention provides a nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of: SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15 and SEQ ID NO:17.

[0020] By a polynucleotide having a nucleotide sequence at least, for example, 95% "identical" to a reference nucleotide sequence is intended that the nucleotide sequence of the polynucleotide is identical to the reference sequence except that the polynucleotide sequence may include up to five point mutations per each 100 nucleotides of the reference nucleotide sequence encoding the pyruvate carboxylase polypeptide. In other words, to obtain a polynucleotide having a nucleotide sequence at least 95% identical to a reference nucleotide sequence, up to 5% of the nucleotides in the reference sequence may be deleted or substituted with another nucleotide, or a number of nucleotides up to 5% of the total nucleotides in the reference sequence may be inserted into the reference sequence.

[0021] As a practical matter, whether any particular nucleic acid molecule is at least 90%, 95%, 97%, 98% or 99% identical to, for instance, the nucleotide sequence shown in Figure 1 or to the nucleotide sequence of the deposited DNA can be determined conventionally using known computer programs such as the FastA program. FastA performs a Pearson and Lipman search for similarity between a query sequence and a group of sequences of the same type nucleic acid. Professor William Pearson of the University of Virginia Department of Biochemistry wrote the FASTA program family (FastA, TFastA, FastX, TFastX

and SSearch). In collaboration with Dr. Pearson, the programs were modified and documented for distribution with GCG Version 6.1 by Mary Schultz and Irv Edelman, and for Versions 8 through 10 by Sue Olson.

[0022] Unless otherwise indicated, all nucleotide sequences determined by sequencing a DNA molecule herein were determined using an automated DNA sequencer (such as the ABI Prism 377). Therefore, as is known in the art for any DNA sequence determined by this automated approach, any nucleotide sequence determined herein may contain some errors. Nucleotide sequences determined by automation are typically at least about 90% identical, more typically at least about 95% to at least about 99.9% identical to the actual nucleotide sequence of the sequenced DNA molecule.

[0023] Unless otherwise indicated, each "nucleotide sequence" set forth herein is presented as a sequence of deoxyribonucleotides (abbreviated A, G , C and T). However, by "nucleotide sequence" of a nucleic acid molecule or polynucleotide is intended, for a DNA molecule or polynucleotide, a sequence of deoxyribonucleotides, and for an RNA molecule or polynucleotide, the corresponding sequence of ribonucleotides (A, G, C and U) where each thymidine deoxynucleotide (T) in the specified deoxynucleotide sequence ~~in~~ is replaced by the ribonucleotide uridine (U). For instance, reference to an RNA molecule having the sequence of SEQ ID NO:1 set forth using deoxyribonucleotide abbreviations is intended to indicate an RNA molecule having a sequence in which each deoxynucleotide A, G or C of SEQ ID NO:1 has been replaced by the corresponding ribonucleotide A, G or C, and each deoxynucleotide T has been replaced by a ribonucleotide U.

[0024] As indicated, nucleic acid molecules of the present invention may be in the form of RNA, such as mRNA, or in the form of DNA, including, for instance, DNA and genomic DNA obtained by cloning or produced synthetically. The DNA may be double-stranded or single-stranded. Single-stranded DNA or RNA

may be the coding strand, also known as the sense strand, or it may be the non-coding strand, also referred to as the anti-sense strand.

[0025] By "isolated" nucleic acid molecule(s) is intended a nucleic acid molecule, DNA or RNA, which has been removed from its native environment. For example, recombinant DNA molecules contained in a vector are considered isolated for the purposes of the present invention. Further examples of isolated DNA molecules include recombinant DNA molecules maintained in heterologous host cells or purified (partially or substantially) DNA molecules in solution. Isolated RNA molecules include *in vivo* or *in vitro* RNA transcripts of the DNA molecules of the present invention. Isolated nucleic acid molecules according to the present invention further include such molecules produced synthetically.

[0026] In another aspect, the invention provides an isolated nucleic acid molecule comprising a polynucleotide which hybridizes under stringent hybridization conditions to a portion of the polynucleotide in a nucleic acid molecule of the invention described herein. By "stringent hybridization conditions" is intended overnight incubation at 42°C in a solution comprising: 50% formamide, 5x SSC (150 mM NaCl, 15mM trisodium citrate), 50 mM sodium phosphate (pH7.6), 5x Denhardt's solution, 10% dextran sulfate, and 20 μg/ml denatured, sheared salmon sperm DNA, followed by washing the filters in 0.1x SSC at about 65°C. By a polynucleotide which hybridizes to a "portion" of a polynucleotide is intended a polynucleotide (either DNA or RNA) hybridizing to at least about 15 nucleotides (nt), and more preferably at least about 20 nt, still more preferably at least about 30 nt, and even more preferably about 30-70 nt of the reference polynucleotide. These are useful as diagnostic probes and primers.

[0027] Of course, polynucleotides hybridizing to a larger portion of the reference polynucleotide (e.g., the deposited plasmid), for instance, a portion 25-750 nt in length, or even to the entire length of the reference polynucleotide, are also useful as probes according to the present invention, as are polynucleotides corresponding to most, if not all, of the nucleotide sequences of any of the

nucleotide sequences included in the present intention. By a portion of a polynucleotide of "at least 20 nt in length," for example, is intended 20 or more contiguous nucleotides from any of the nucleotide sequences of the reference polynucleotides, (e.g., the deposited DNA or the nucleotide sequence as shown in any of the figures). As indicated, such portions are useful diagnostically either as a probe, according to conventional DNA hybridization techniques, or as primers for amplification of a target sequence by the polymerase chain reaction (PCR), as described, for instance, in *Molecular Cloning, A Laboratory Manual*, 2nd. edition, edited by Sambrook, J., Fritsch, E. F. and Maniatis, T., (1989), Cold Spring Harbor Laboratory Press, the entire disclosure of which is hereby incorporated herein by reference.

[0028] The nucleic acid molecules of the present invention are suitable for use in vectors. As such, polynucleotides of interest can be joined to the nucleic acid molecules of the present invention, which may optionally contain selectable markers. A preferred embodiment of the present invention is that the vector comprises a functional *Corynebacterium* replication origin. A replication origin is a nucleotide sequence, typically several hundred base pairs long, that is vital to the initiation of DNA replication.

[0029] The vectors can optionally contain an exogenous terminator of transcription; an exogenous promoter; and a discrete series of restriction endonuclease recognition sites, said series being between said promoter and said terminator. The vector can optionally contain their native expression vectors and/or expression vectors which include chromosomal-, and episomal-derived vectors, e.g., vectors derived from bacterial exogenous plasmids, bacteriophage, and vectors derived from combinations thereof, such as cosmids and phagemids.

[0030] A DNA insert of interest should be operatively linked to an appropriate promoter, such as its native promoter or a host-derived promoter, the phage lambda P_L promoter, the phage lambda P_R promoter, the *E. coli lac* promoters, such as the *lac*I and *lac*Z promoters, *trp* and *tac* promoters, the T3 and T7

promoters and the *gpt* promoter to name a few. Other suitable promoters will be known to the skilled artisan.

[0031] The expression constructs will further contain sites for transcription initiation, termination and, in the transcribed region, a ribosome binding site for translation. The coding portion of the mature transcripts expressed by the constructs can include a translation initiating codon at the beginning and a termination codon appropriately positioned at the end of the polypeptide to be translated.

[0032] As indicated, the expression vectors will preferably include at least one selectable marker. Preferably the selection marker comprises a nucleotide sequence which confers antibiotic resistance in a host cell population. Such markers include amikacin, augmentin (amoxicillin plus clavulonic acid), ampicillin, cefazolin, cefoxitin, ceftazidime, ceftiofur, cephalothin, enrofloxacin, florfenicol, gentamicin, imipenem, kanamycin, penicillin, sarafloxicin, spectinomycin, streptomycin, tetracycline, ticarcillin, tilmicosin, or chloramphenicol resistance genes. Other suitable markers will be readily apparent to the skilled artisan.

[0033] The invention also provides for a method of producing a host cell where the expression vectors of the current invention have been introduced into the host cell. Methods of introducing genetic material into host cells, such as those described in typical molecular biology laboratory manuals, for example J. Sambrook, E.F. Fritsch and T. Maniatis, *Molecular Cloning: A Laboratory Manual*, 2d ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York (1989), are well known to the skilled artisan. These methods include, but are not limited to, calcium phosphate transfection, DEAE-dextran mediated transfection, microinjection, lipid-mediated transfection, electroporation or infection. Accordingly, a preferred embodiment of the present invention provides a host cell comprising the vector of the present invention.

[0034] As used in the present invention, a host cell refers to any prokaryotic or eukaryotic cell where the desired nucleic acid sequence has been introduced into the cell. There are a variety of suitable host cells, including but not limited to bacterial, fungal, insect, mammalian and plant cells, that can be utilized in the present invention. Representative bacterial host cells include, but are not limited to, *Streptococci, Staphylococci, E. coli, Streptomyces, Bacillus* and *Corynebacterium.* Representative fungal cells include but are not limited to, yeast cells and *Aspergillus.* Insect cells include, but are not limited to, *Drosophila* S2 and *Spodoptera* Sf9 cells. Examples of mammalian cells include, but are not limited to, CHO, COS and Hela cells.

[0035] The present invention provides methods for utilizing the nucleic acid of SEQ ID NO:1 or SEQ ID NO:3, which encodes the amino acid sequence of a mutant pyruvate carboxylase. Such methods include the replacement of the wild-type pyruvate carboxylase with the feedback-resistant pyruvate carboxylase, and the production of amino acids. The method for replacement of a wild-type pyruvate carboxylase gene, with a feedback resistant pyruvate carboxylase gene, in a *Corynebacterium glutamicum* host cell comprises the steps of: (a) replacing a genomic copy of the wild-type pyruvate carboxylase gene with a selectable marker gene through homologous recombination to form a first recombinant strain; and (b) replacing the selectable marker gene of step (a) in the first recombinant strain, with the feedback resistant pyruvate carboxylase gene through homologous recombination to form a second recombinant strain. The homologous recombination in steps (a) and (b) would occur between the genetic material of the host cell and any of the vectors of the present invention.

[0036] Homologous recombination is a technique that is used to disrupt endogenous nucleotide sequences in a host cell. Normally, when an exogenous nucleotide sequence is inserted into a host cell, this polynucleotide may randomly insert into any area of the host cell's genome, including endogenous plasmids. However, with homologous recombination, the exogenous nucleotide

sequence contains sequences that are homologous to an endogenous nucleotide sequence within the host cell. Once introduced into the cell, for example by electroporation, the exogenous nucleotide sequence will preferentially recombine with and replace the endogenous nucleotide sequence with which it is homologous.

[0037] As used herein, an exogenous nucleotide sequence, is a nucleotide sequence which is not found in the host cell. Thus, the term exogenous nucleotide sequence is meant to encompass a nucleotide sequence that is foreign to the host cell, as well as a nucleotide sequence endogenous, or native, to the host cell that has been modified. Modification of the endogenous nucleotide sequence may include, for instance, mutation of the native nucleotide sequence or any of its regulatory elements. As used herein, mutation is defined as any change in the wild-type sequence of the host's genetic material, including plasmid DNA. An additional form of modification may also include fusion of the endogenous nucleotide sequence to a nucleotide sequence that is normally not present, in relation to the endogenous nucleotide sequence.

[0038] Host cells that have undergone homologous recombination are selected on the basis of antibiotic resistance through the use of, for example, the selectable markers mentioned above. The process of selecting cells that have undergone homologous recombination will be readily apparent to one skilled in the art.

[0039] Another aspect of the current invention is a method for producing amino acids. In the current context, production of amino acids is accomplished by culturing host cells where a vector of the present invention has been introduced into the host cell, or culturing host cells where homologous recombination, involving a vector of the present invention, has taken place. Culturing of the host cells is performed in the appropriate culture media. Subsequent to culturing the host cells in culture media, the desired amino acids are separated from the culture media. Preferably, the amino acids produced by the methods described herein include L-lysine, L-threonine, L-methionine, L-isoleucine, L-glutamate, L-

arginine and L-proline. More preferably, the present invention relates to the production of L-lysine.

[0040] The present invention provides an isolated or purified polypeptide encoded by the DNA plasmid encoding pyruvate carboxylase contained in Deposit Number NRRL B-30293, the amino acid sequence of SEQ ID NO:2 or the amino acid sequence of SEQ ID NO:4. Still another aspect of the present invention provides a polypeptide comprising the amino acid sequence selected from the group consisting of SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14, SEQ ID NO:16 and SEQ ID NO:18.

[0041] Accordingly, SEQ ID NO:6 corresponds to the amino acid sequence: PSKNIDDIVKSAE. SEQ IN NO:8 corresponds to the amino acid sequence: RGMRFVSSPDELR. SEQ ID NO:10 corresponds to the amino acid sequence: AAFGDGSVYVERA. SEQ ID NO:12 corresponds to the amino acid sequence: VQILGDRTGEVVH. SEQ ID NO:14 corresponds to the amino acid sequence: IATGFIGDHPHLL. SEQ ID NO:16 corresponds to the amino acid sequence: TITASVEGKIDRV. SEQ ID NO:18 corresponds to the amino acid sequence: MTAITLGGLLLKGIITLV.

[0042] All of the polypeptides of the present invention are preferably provided in an isolated form. As used herein, "isolated polypeptide" is intended to mean a polypeptide removed from its native environment. Thus, a polypeptide produced and/or contained within a recombinant host cell is considered isolated for purposes of the present invention. Also intended as an "isolated polypeptide" are polypeptides that have been purified, partially or substantially, from a recombinant host. For example, a recombinantly produced version of the pyruvate carboxylase enzyme can be substantially purified by the one-step method described in Smith and Johnson, *Gene 67*:31-40 (1988).

[0043] One aspect of the present invention include the polypeptides which are at least 80% identical, more preferably at least 90%, 95% or 100% identical to the polypeptide encoded by the DNA plasmid encoding pyruvate carboxylase

contained in Deposit Number NRRL B-30293, the polypeptide of SEQ ID NO:2 or the polypeptide of SEQ ID NO:4.

[0044] By a polypeptide having an amino acid sequence at least, for example, 95% "identical" to the amino acid sequence of SEQ ID NO:2, for example, it is intended that the amino acid sequence of the polypeptide is identical to the reference sequence except that the polypeptide sequence may include up to five amino acid alterations per each 100 amino acids of the amino acid sequence of SEQ ID NO:2, for example. In other words, to obtain a polypeptide having an amino acid sequence at least 95% identical to a reference amino acid sequence, up to 5% of the amino acid residues in the reference sequence may be deleted or substituted with another amino acid, or a number of amino acids up to 5% of the total amino acid residues in the reference sequence may be inserted into the reference sequence. These alterations of the reference sequence may occur at the amino or carboxy terminal positions of the reference amino acid sequence or anywhere between those terminal positions, interspersed either individually among residues in the reference sequence or in one or more contiguous groups within the reference sequence.

[0045] As a practical matter, whether any particular polypeptide is, for instance, 95% identical to the amino acid sequence shown in SEQ ID NO:2, SEQ ID NO:4 or to the amino acid sequence encoded by deposited DNA clone can be determined conventionally using known computer programs such the Bestfit program (Wisconsin Sequence Analysis Package, Version 8 for Unix, Genetics Computer Group, University Research Park, 575 Science Drive, Madison, WI 53711). When using Bestfit or any other sequence alignment program to determine whether a particular sequence is, for instance, 95% identical to a reference sequence according to the present invention, the parameters are set, of course, such that the percentage of identity is calculated over the full length of the reference amino acid sequence and that gaps in homology of up to 5% of the total number of amino acid residues in the reference sequence are allowed.

[0046] Another aspect of the present invention provides a nucleic acid molecule encoding the polypeptide comprising the amino acid sequence selected from the group consisting of SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14, SEQ ID NO:16 and SEQ ID NO:18. Preferably, the invention provides for nucleic acid molecules, which code for the aforementioned polypeptides, that are selected from the group consisting of SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15 and SEQ ID NO:17.

[0047] Accordingly, SEQ ID NO:5 corresponds to the nucleic acid sequence that codes for the amino acid sequence of SEQ ID NO:6. SEQ ID NO:7 corresponds to the nucleic acid sequence that codes for the amino acid sequence of SEQ ID NO:8. SEQ ID NO:9 corresponds to the nucleic acid sequence that codes for the amino acid sequence of SEQ ID NO:10. SEQ ID NO:11 corresponds to the nucleic acid sequence that codes for the amino acid sequence of SEQ ID NO:12. SEQ ID NO:13 corresponds to the nucleic acid sequence that codes for the amino acid sequence of SEQ ID NO:14. SEQ ID NO:15 corresponds to the nucleic acid sequence that codes for the amino acid sequence of SEQ ID NO:16. SEQ ID NO:17 corresponds to the nucleic acid sequence that codes for the amino acid sequence of SEQ ID NO:18.

[0048] Methods used and described herein are well known in the art and are more particularly described, for example, in J.H. Miller, *Experiments in Molecular Genetics*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York (1972); J.H. Miller, *A Short Course in Bacterial Genetics*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York (1992); M. Singer and P. Berg, *Genes & Genomes*, University Science Books, Mill Valley, California (1991); J. Sambrook, E.F. Fritsch and T. Maniatis, *Molecular Cloning: A Laboratory Manual*, 2d ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York (1989); P.B. Kaufman *et al.*, *Handbook of Molecular and Cellular Methods in Biology and Medicine*, CRC Press, Boca Raton, Florida (1995); *Methods in*

Plant Molecular Biology and Biotechnology, B.R. Glick and J.E. Thompson, eds., CRC Press, Boca Raton, Florida (1993); P.F. Smith-Keary, *Molecular Genetics of Escherichia coli*, The Guilford Press, New York, NY (1989); *Plasmids: A Practical Approach*, 2nd Edition, Hardy, K.D., ed., Oxford University Press, New York, NY (1993); *Vectors: Essential Data*, Gacesa, P., and Ramji, D.P., eds., John Wiley & Sons Pub., New York, NY (1994); *Guide to Electroporation and electrofusions*, Chang, D., *et al.*, eds., Academic Press, San Diego, CA (1992); *Promiscuous Plasmids of Gram-Negative Bacteria*, Thomas, C.M., ed., Academic Press, London (1989); *The Biology of Plasmids*, Summers, D.K., Blackwell Science, Cambridge, MA (1996); *Understanding DNA and Gene Cloning: A Guide for the Curious*, Drlica, K., ed., John Wiley and Sons Pub., New York, NY (1997); *Vectors: A Survey of Molecular Cloning Vectors and Their Uses*, Rodriguez, R.L., *et al.*, eds., Butterworth, Boston, MA (1988); *Bacterial Conjugation*, Clewell, D.B., ed., Plenum Press, New York, NY (1993); Del Solar, G., *et al.*, Replication and control of circular bacterial plasmids," *Microbiol. Mol. Biol. Rev. 62*:434-464 (1998); Meijer, W.J., *et al.*, "Rolling-circle plasmids from Bacillus subtilis: complete nucleotide sequences and analyses of genes of pTA1015, pTA1040, pTA1050 and pTA1060, and comparisons with related plasmids from gram-positive bacteria," *FEMS Microbiol. Rev. 21*:337-368 (1998); Khan, S.A., "Rolling-circle replication of bacterial plasmids," *Microbiol. Mol. Biol. Rev. 61*:442-455 (1997); Baker, R.L., "Protein expression using ubiquitin fusion and cleavage," *Curr. Opin. Biotechnol. 7*:541-546 (1996); Makrides, S.C., "Strategies for achieving high-level expression of genes in *Escherichia coli*," *Microbiol. Rev. 60*:512-538 (1996); Alonso, J.C., *et al.*, "Site-specific recombination in gram-positive theta-replicating plasmids," *FEMS Microbiol. Lett. 142*:1-10 (1996); Miroux, B., *et al.*, "Over-production of protein in *Escherichia coli*: mutant hosts that allow synthesis of some membrane protein and globular protein at high levels," *J. Mol. Biol. 260*:289-298 (1996); Kurland, C.G., and Dong, H., "Bacterial growth inhibited by overproduction of protein,"

Mol. Microbiol. 21:1-4 (1996); Saki, H., and Komano, T., "DNA replication of IncQ broad-host-range plasmids in gram-negative bacteria," *Biosci. Biotechnol. Biochem. 60*:377-382 (1996); Deb, J.K., and Nath, N., "Plasmids of corynebacteria," *FEMS Microbiol. Lett. 175*:11-20 (1999); Smith, G.P., "Filamentous phages as cloning vectors," *Biotechnol. 10*:61-83 (1988); Espinosa, M., *et al.*, "Plasmid rolling cicle replication and its control," *FEMS Microbiol. Lett. 130*:111-120 (1995); Lanka, E., and Wilkins, B.M., "DNA processing reaction in bacterial conjugation," *Ann. Rev. Biochem. 64*:141-169 (1995); Dreiseikelmann, B., "Translocation of DNA across bacterial membranes," *Microbiol. Rev. 58*:293-316 (1994); Nordstrom, K., and Wagner, E.G., "Kinetic aspects of control of plasmid replication by antisense RNA," *Trends Biochem. Sci. 19*:294-300 (1994); Frost, L.S., *et al.*, "Analysis of the sequence gene products of the transfer region of the F sex factor," *Microbiol. Rev. 58*:162-210 (1994); Drury, L., "Transformation of bacteria by electroporation," *Methods Mol. Biol. 58*:249-256 (1996); Dower, W.J., "Electroporation of bacteria: a general approach to genetic transformation," *Genet. Eng. 12*:275-295 (1990); Na, S., *et al.*, "The factors affecting transformation efficiency of coryneform bacteria by electroporation," *Chin. J. Biotechnol. 11*:193-198 (1995); Pansegrau, W., "Covalent association of the traI gene product of plasmid RP4 with the 5'-terminal nucleotide at the relaxation nick site," *J. Biol. Chem. 265*:10637-10644 (1990); and Bailey, J.E., "Host-vector interactions in *Escherichia coli*," *Adv. Biochem. Eng. Biotechnol. 48*:29-52 (1993).

EXAMPLES

[0049] The following examples are illustrative only and are not intended to limit the scope of the invention as defined by the appended claims.

Strains and Media

[0050] Bacterial strains used were *Corynebacterium glutamicum* ATCC 21253 and NRRL B-11474. These strains have an auxotrophy for homoserine (ATCC 21253) and for threonine, methionine and alanine (NRRL B-11474).

[0051] Defined medium for *Corynebacterium glutamicum* ATCC 21253 contained the following ingredients (per liter): glucose, 20 g; NaCl, 2 g; citrate (trisodium salt, dihydrate), 3 g; $CaCl_2 \cdot 2H_2O$, 0.1 g; $MgSO_4 \cdot 7H_2O$, 0.5 g; $Na_2EDTA \cdot 2H_2O$, 75 mg; $FeSO_4 \cdot 7H_2O$, 50 mg; 100x salt solution, 20 ml; K_2HPO_4, 4 g; KH_2PO_4, 2 g; $(NH_4)_2SO_4$, 7.5 g; urea, 3.75 g; leucine, 0.1 g; threonine, 0.15 g; methionine, 0.05 g; thiamine, 0.45 mg; biotin, 0.45 mg; pantothenic acid, 4.5 mg (pH 7.0). The salt solution contained the following ingredients (per liter): $MnSO_4$, 200 mg; $Na_2B_4O_7 \cdot 10H_2O$, 20 mg; $(NH_4)_6Mo_7O_{24} \cdot 4H_2O$, 10 mg; $FeCl_3 \cdot 6H_2O$, 200 mg; $ZnSO_4 \cdot 7H_2O$, 50 mg; $CuCl_2 \cdot 2H_2O$, 20 mg (pH 2.0).

[0052] Defined medium for *Corynebacterium glutamicum* NRRL B-11474 contained the following ingredients (per liter): glucose, 20 g; NaCl, 1 g, $MgSO_4 \cdot 7H_2O$, 0.4 g; $FeSO_4 \cdot 7H_2O$, 0.01 g; $MnSO_4 \cdot H_2O$, 0.01 g; KH_2PO_4, 1 g; $(NH_4)_2SO_4$, 10 g; urea, 2.5 g; alanine, 0.5 g; threonine, 0.25 g; methionine, 0.5 g; thiamine, 0.45 mg; biotin, 0.45 mg; niacinamide, 50 mg (pH 7.2).

[0053] *Pyruvate Carboxylase and Phosphoenol Pyruvate Carboxylase Assay*

[0054] Pyruvate carboxylate and phosphoenol pyruvate carboxylate assays were performed with permeabilized cells prepared by the following method. Log phase cells were harvested by centrifugation for 10 min at 5000 xg at 4°C and washed

with 20 ml of the ice-cold washing buffer (50 mM Tris/HCl [pH 6.3] containing 50 mM NaCl). The cell pellet was resuspended in an ice-cold Hepes buffer (100 mM Hepes [pH 7.5] containing 20% Glycerol) to reach a final concentration of 25 g dry cell weight/liter. Resuspended cells were permeabilized by adding 30 µl of a 10% Hexadecyltrimethyl-ammonium bromide (CTAB) (w/v) solution to 1 ml of cells to give a final concentration of 0.3% (CTAB)(v/v).

[0055] For determination of pyruvate carboxylate activity, the assay mixture contained 10 mM pyruvic acid, 14 mM $KHCO_3$, 4 mM $MgCl_2$, 1.75 mM ATP, 50 µmole acetyl-CoA, 0.3 mg bovine serum albumin, 0.055 U citrate synthase and 50 mM sodium phosphate buffer ([pH 7.5] containing 0.1 mg 5, 5′-Dithio-bis(2-nitrobenzoic acid) (DTNB)) in a final volume of 1 ml. The reaction was started at 30°C with the addition of 10 µl of the permeabilized cell suspension, and the formation of DTNB-thiophenolate was followed over time at 412 nm. Relevant standards and controls were carried out in the same manner.

[0056] For determination of phosphoenol pyruvate carboxylase activity, the assay mixture contained 10 mM phosphoenol pyruvate, 14 mM $KHCO_3$, 4 mM $MgCl_2$, 50 µmole acetyl-CoA, 0.3 mg bovine serum albumin, 0.055 U citrate synthase and 50 mM sodium phosphate buffer ([pH 7.5] containing 0.1 mg 5, 5′-Dithio-bis(2-nitrobenzoic acid) (DTNB)) in a final volume of 1 ml. The reaction was carried out in the same conditions described for the pyruvate carboxylase assay.

[0057] The reproducibility for enzyme assays was typically 10%.

DNA Isolation and Purification

[0058] DNA was isolated from cultures of NRRL B-11474 cells. Defined media for NRRL B-11474 (CM media) contain the following ingredients, per liter: sucrose, 50 g; KH_2PO_4, 0.5 g; K_2HPO_4, 1.5 g; urea, 3 g; $MgSO_4 \cdot 7H_2O$, 0.5 g; polypeptone, 20 g; beef extract, 5 g; biotin, 12.5 ml (60 mg/L); thiamine, 25 ml (120 mg/L), niacinamide, 25 ml (5g/L); L-methionine, 0.5 g; L-threonine, 0.25 g;

L-alanine, 0.5 g. NRRL B-11474 cells were harvested from CM media and suspended in 10 ml of TE, pH 8 (10 mM Tris*Cl, 1 mM EDTA). Forty micrograms of RNase A and 10 milligrams of lysozyme were added per milliliter of suspension and the suspension was incubated at 37°C for 30 minutes. The suspension was made in 1.0% in sodiumdodecyl sulfate (SDS) and 0.1 mg/l proteinase K was added, and the cells were lysed by incubation at 37°C for 10 minutes. Nucleic acids were purified by three extractions with TE-saturated phenol (pH7), followed by ethanol precipitation. Nucleic acid precipitates were twice washed with 80% ethanol and redissolved in TE pH 8.

[0059] The concentrations of DNA were quantified spectrophotometrically at 260 nm. Purity of DNA preparations were determined spectrophotometrically (A260/A280 and A26JA230 ratios) and by agarose gel electrophoresis (0.8% agarose in 1x TAE).

[0060] Sequencing of the genomic DNA was performed, as is known by one of ordinary skill in the art, by creating libraries of plasmids and cosmids using pGEM3 and Lorist 6 respectively. Briefly, a Sau3AI digestion was performed on the genomic DNA and inserted into the BamHI site of pGEM3. The forward primer was used to generate a sequence, and primer walking generated the remainder of the sequence.

Activity of Pyruvate Carboxylase

Development of a Continuous Assay for Determining Pyruvate Carboxylase Activity

[0061] A discontinuous assay for determining pyruvate carboxylase from permeabilized cells has been previously described (Peters-Wendisch, P.G. *et al. Microbiology,* 143: 1095-1103 (1997)). Because of the central location of OAA in the metabolism, it seemed to be that OAA would accumulate during the first reaction of the discontinuous assay. Most likely, OAA would be lost to other

products, because of the competing enzymes that are still active. This depletion of OAA would inevitably lead to the underestimation of pyruvate carboxylase activity. To verify this assumption of decreasing OAA concentrations, a known amount of OAA was added to the first reaction in presence of permeabilized and non-permeabilized cells. A significant loss of OAA was detected, demonstrating that permeabilized cells are capable of further transformation of OAA.

[0062] To account for the intrinsic loss of OAA during the experiment, a continuous assay was carried out by coupling the two-reaction assay to a one-reaction assay in presence of an excess of citrate synthase. The amount of permeabilized cells added in the assay was optimized to obtain a detectable activity, with the lowest possible background absorbency due to the presence of cells.

[0063] To confirm that the continuous assay specifically detected pyruvate carboxylase activity, controls were carried out by assaying for activity in absence of each reaction component (Table 1). Using these controls, the detected activity was determined to be a carboxylation reaction requiring pyruvate, Mg and ATP.

Table 1: Controls for the continuous pyruvate carboxylate assay.

Control	Detected Activity (Abs/min.mg DCW)
Complete mixture	0.30
Cells omitted	0
Pyruvate omitted	0.01
$KHCO_3$ omitted	0.03
$MgCl_2$ omitted	0.02
ATP omitted	0.03
Citrate synthase omitted	0.10
Complete + biotin	0.35
Complete + avidin	Not determined yet

To optimize the assay, the influence of the ratio of CTAB:cells was tested. Maximal activity was measured between 8 and 24 mg CTAB/mg dry cell weight (DCW). Pyruvate carboxylase activity was measured in cells incubated with

CTAB with varying incubation times. The activity of pyruvate carboxylase remained constant within 0 and 5 minutes. Similarly, different concentrations of DTNB, within the range 0.1-0.3 g/l, gave identical pyruvate carboxylase activity. To confirm the ability of the assay for determining pyruvate carboxylase activity in *Corynebacterium glutamicum*, different quantities of cells were used. Linearity between enzyme activity and quantity of cells was observed within the range 0-0.3 mg DCW.

Enzymology Study of Pyruvate Carboxylase from Corynebacterium glutamicum: Behavior of Pyruvate Carboxylase Towards Its Substrates

[0064] Pyruvate carboxylase activity was determined as a function of various concentrations of its substrates: pyruvate, bicarbonate and ATP (Figure 4). Based on the data generated, the affinity constants of pyruvate carboxylase for its substrates were determined (Table 2). The pyruvate carboxylase from NRRL B-11474 and ATCC 21253 strains demonstrated a similar affinity for pyruvate and ATP. Pyruvate carboxylase activity in both strains were inhibited by ATP above a concentration of 2 mM. However pyruvate carboxylase in ATCC 21253 had a higher affinity for bicarbonate than pyruvate carboxylase from NRRL B-11474.

Strain	$K_{M(pyruvate)}$[mM]	$K_{M(HCO_3^-)}$[mM]	$K_{M(ATP)}$[mM]
C. glutamicum			
Pyc BF100	1.3 ± 03	14.4 ± 4	0.4 ± 0.1
Pyc ATCC 21253	0.3 ± 0.1	2.9 ± 0.8	0.3 ± 0.1

Table 2: Comparison of affinity constants for substrates on pyruvate carboxylate from *C. glutamicum* BF100 and ATCC 21253.

Aspartate Inhibition of Pyruvate Carboxylase

[0065] Aspartate inhibits phosphoenol pyruvate carboxylase (PEPC) activity. To determine the effect of aspartate on the activity of pyruvate carboxylase, aspartate was added at different concentrations in the spectrophotometer cuvette and enzyme activities were measured. As a comparison, the same experiment was carried out with PEPC in ATCC 21253 (Figure 5).

[0066] The PEPC of *Corynebacterium glutamicum* (ATCC 21253) was found to be strongly inhibited by aspartate. The enzyme was completely inhibited with a concentration of 5 mM aspartate. However, pyruvate carboxylase from the same strain was less sensitive to aspartate, i.e. it retained 35% of its original activity in the presence of 25 mM aspartate.

[0067] The pyruvate carboxylase activity in NRRL B-11474 showed a higher basal pyruvate carboxylase activity than ATCC 21253, i.e. the pyruvate carboxylase activity was about 5-times higher in NRRL B-11474 than in the ATCC 21253. Moreover, a dramatic difference in their aspartate inhibition patterns was found. Pyruvate carboxylase from NRRL B-11474 strain was activated by low aspartate concentrations within the range 0-30 mM and inhibited within the range 30-100 mM aspartate. Nevertheless it retained 50% of its original activity, even in the presence of 100 mM aspartate. Activity was maintained at 30% in the presence of 500 mM aspartate. On the other hand, Pyruvate carboxylase from ATCC 21253 was found to be more sensitive to aspartate than pyruvate carboxylase from NRRL B-11474. The pyruvate carboxylase from ATCC 21253 lost 70% of its original activity at a concentration of 30 mM aspartate.

Activation of Pyruvate Carboxylase by Acetyl-CoA

add B13

[0068] Pyruvate carboxylase activity was measured in the presence of different concentrations of acetyl-CoA (Figure 6). Pyruvate carboxylase activity in both strains increased with increasing acetyl-CoA concentrations. The effect of acetyl-CoA on citrate synthase itself was studied also. Acetyl-CoA had a Km of 10 μM, demonstrating that under our conditions, citrate synthase is saturated with acetyl-CoA. Therefore, the increasing activity of pyruvate carboxylase with increasing acetyl-CoA concentration is the result of acetyl-CoA activating pyruvate carboxylase.

* * * * *

[0069] All publications mentioned herein above are hereby incorporated in their entirety by reference.

[0070] While the foregoing invention has been described in some detail for purposes of clarity and understanding, it will be appreciated by one skilled in the art from a reading of this disclosure that various changes in form and detail can be made without departing from the true scope of the invention and appended claims.